

Inconsistencies in the application of harmonic analysis to pulsating stars

J. Pascual-Granado¹, R. Garrido¹, and J. C. Suárez^{2, 1}

¹Instituto de Astrofísica de Andalucía (CSIC), Granada 18008, Spain

²Dept. Física Teórica y del Cosmos. Universidad de Granada. 18071 Granada. Spain
e-mail: javier@iaa.es, garrido@iaa.es, jcsuarez@ugr.es

July 29, 2015

ABSTRACT

Using ultra-precise data from space instrumentation we found that the underlying functions of stellar light curves from some AF pulsating stars are non-analytic, and consequently their Fourier expansion is not guaranteed. This result demonstrates that periodograms do not provide a mathematically consistent estimator of the frequency content for this kind of variable stars. More importantly, this constitutes the first counterexample against the current paradigm which considers that any physical process is described by a continuous (band-limited) function that is infinitely differentiable.

Key words. Asteroseismology – Methods: data analysis – Stars: oscillations – Kepler -

1. Introduction

The unexpected huge number of frequencies found in multi periodic stars (Poretti 2009) and the ubiquitous presence of correlated noise in the residuals of the fitting of the light curves of pulsating stars, could have a common origin.

The necessary condition for a correct Fourier analysis, i.e. to have a Fourier expansion, is guaranteed when the function is analytic, otherwise Fourier analysis is not a consistent approximation. In this work we examine the differentiability of the function describing photometric data of pulsating stars obtained by CoRoT (Auvergne et al. 2009) and *Kepler* (Gilliland et al. 2010). We determine how smoothly the data points of the light curve can be fitted in order to fully reproduce the function. We call this fine structure property the connectivity of the function at a given point. We studied this property by means of two numerical approaches: cubic splines - analytic¹, and autoregressive moving average (Box 1976) methods - non-analytic.

2. Analyticity

Fourier analysis is a useful technique for frequency detection. In order to apply Fourier techniques to a given function, Parseval's theorem (Parseval 1992) must hold. This theorem states that the integral of the squared modulus of a function is equal to the integral of the squared modulus of its Fourier transform. For this to be true the function must be square integrable. The Fourier expansion of the function converges pointwise almost everywhere when it is square integrable (Carleson 1966). Unlike the Fourier series, which is an infinite sum, the Discrete Fourier Transform (DFT) and its inverse, the IDFT, which is the analogous to the Fourier series, always converge if the sequence is finite, because

they are finite sums. But when the Fourier series of the function is divergent, or not convergent to the value of the function at this point, the DFT no longer provides a mathematical description based on the frequency components of the function. That is, a frequency is only a well-defined physical variable when the Fourier series converges. Therefore, before a Fourier analysis can be applied to a time series, the convergence of the Fourier expansion of the function must be demonstrated.

On the other hand, the analyticity of a function is a sufficient condition for the square integrability (Dijk 2009). A function is said to be analytic if it is infinitely differentiable so that its Taylor expansion is convergent. If this condition is not met the convergence of the Fourier series is not guaranteed. That is, analyticity implies that Dirichlet conditions for convergence are fulfilled.

In asteroseismology studies (and often in astrophysics) such a convergence of the Fourier series is assumed, i.e. the analyticity of the underlying function is thus also assumed. To date, no methods to verify it have been proposed. Here we propose a practical method to verify the analyticity condition of the underlying function from the observed sampled data, which in consequence provide evidence of the convergence of the Fourier series.

In this sense, analyticity refers in this paper to the property that a given function can be expanded in Fourier series. To study the analyticity of a function we introduce a numerical approximation for the differentiability of discrete time series: the connectivity².

3. Connectivity

Time-dependent astronomical data are given as a discrete sequence of samples from an underlying function, which is a representation of a given physical observable. This latter is under-

¹ By definition only the first and second derivatives are continuous for cubic splines but we are interested here only in the first derivative. In any case, if the first derivative is not continuous the functions are non-analytic.

² Avoid confusion with graph theory term connectivity (Diestel 2005)

stood as the physical quantity that satisfies all the properties³ derived from the Lebesgue measure definition (Jech 1997).

The discrete sampled sequence of photometric variations of luminosity and/or radial velocity measurements of pulsating stars is commonly called light curve (or RV). Let us consider here such a series as a finitely close sequence of data points which samples a function composed by a set of infinitely close data points. Under these conditions, we want to analyse whether such a light curve fully and unambiguously determines the properties of the sampled function. This is the well-known problem of sampling. It is known (Nyquist-Shannon theorem, Shannon 1949) that this is possible when the sampled function is band-limited to less than half the sampling rate, i.e. the Nyquist frequency. When this happens the complete information of the continuum (i.e. the sampled function) is contained in the discrete sequence. This implies that all the properties of the function can be studied, even those concerning the pointwise limits of the function. Consequently, we are allowed to adequately evaluate the differentiability of the sampled function.

Let us now generalize our discrete sequence of data to sample a function composed by an stochastic sequence plus a deterministic contribution that can have or not a wave-like structure. Such decomposition is guaranteed under the hypothesis of stationarity by the Wold theorem (Wold 1938). The random component might complicate the differentiability analysis of a time series, however the Kolmogorov continuity theorem allow us to fully characterize the properties of the function (Revuz 1999) in a similar way as it is done for pure deterministic signals. This requires necessarily the topological separability of the function, which is applicable under certain technical constraints related with the moments of the data differences, particularly variations at most as those given by a power law (see Appendix A for a discussion on these constraints). When these restrictions are fulfilled Kolmogorov theorem states that a stochastic process has a continuous representation.

This is a necessary and sufficient condition for the connectivity analysis that we will define in the next section.

3.1. Definition

Given a data point from a discrete sequence, we can calculate a forward and backward extrapolation from the subsets bracketing the selected data point. Then, it is possible to check whether both extrapolations converge to the same value and coincide with the value of the selected data point.

The above theoretical framework permits the following self-consistent definition of the connectivity C_n of a data point y_n of a discrete sequence which samples the function $F(t)$

$$C_n = \epsilon_n^f - \epsilon_n^b, \quad (1)$$

where ϵ_n^f , ϵ_n^b , are the deviations of the forward and backward extrapolations from the sample y_n

$$\begin{aligned} \epsilon_n^f &= y_n^f - y_n \\ \epsilon_n^b &= y_n^b - y_n \end{aligned} \quad (2)$$

represented by y_n^f and y_n^b , respectively. In fact, a discrete approximation of the derivability condition of F at the point t_n can be:

$$\frac{y_n^b - y_{n+1}}{\Delta t} = \frac{y_n^f - y_{n-1}}{\Delta t} \quad (3)$$

³ We want to stress here that an observable of a given physical property is not necessarily coincident with the function which represents it.

where Δt is the sampling rate of the sequence. Note that connectivity is closely related to derivability. In fact, we can define a new quantity \mathcal{D}_n as:

$$\mathcal{D}_n = \frac{\xi_n + y_{n+1} - y_{n-1}}{2\Delta t} \quad (4)$$

which reduces to the typical point derivative (Burden 2011) for discrete data when ξ_n , the non-differentiable component of the sequence, is zero. Likewise, the connectivity as defined above is a function of ξ_n , that is:

$$C_n = C(\xi_n) + \epsilon_n \quad (5)$$

From this equation, when connectivities are not zero but an independent normal stochastic sequence⁴ they can be identified with the random term ϵ_n which is a numerical error. In this case the non-differentiable component is zero ($\xi_n = 0$) and therefore the derivatives are still well defined. When connectivities are correlated with the signal the non-differentiable component is different from zero ($\xi_n \neq 0$) and therefore a derivative cannot be defined. In this case, the function is non-analytic and it does not satisfy the conditions under which Parseval's theorem is demonstrated, meaning that Fourier analysis is not consistent.

In summary, sampled data are obtained through the sampling of functions that needs not be necessarily differentiable a priori. Connectivities can be interpreted as jumps in the derivatives giving rise to discontinuities. This makes connectivities mathematically self-consistent. They can be considered as equivalent to the non-differentiability coefficient introduced by Wiener (1923) in a different physical context, and allowing to check the differentiability of the underlying functions.

3.2. Method

The forward and backward extrapolations necessary to obtain the connectivities have to be calculated numerically. Two conceptually different numerical approaches were considered. A comparison of both approaches allows us to confirm the results without ambiguity.

3.2.1. Cubic splines

The Stone-Weierstrass theorem (Royden 1988) states that if a function is uniformly continuous in a closed interval it can be approximated arbitrarily well, i.e. as closely as desired, by a polynomial of degree n , n being a natural number. Then, the entire function can be represented by a piecewise polynomial function of finite degree. But piecewise polynomial functions have a drawback, the Runge's phenomenon, showing oscillations at the edges of the approximated function. This is equivalent to the Gibbs phenomenon in the Fourier approximation.

Taking all of this into account, we selected piecewise smooth polynomials for our modelling.

Formulated in this general way any particular property can be extended to any analytic function. That is, if a continuous function can be fitted arbitrarily well with a given cubic spline parametric model then the analyticity condition is automatically satisfied, otherwise the convergence of the Fourier series is not guaranteed for that particular function. In short, the cubic splines approach gives a key information about local features of the function that generates the observed discrete sequence.

⁴ This is the theoretically expected distribution of independent samples. It is usually understood as gaussian white noise.

To calculate the coefficients of the cubic spline functions a tri-diagonal linear system is solved (Boor 1978). The only parameter affecting the spline interpolation that can be adjusted is the number of datapoints of the fitted segments.

3.2.2. ARMA

These models (Box 1976) give a parametric representation of a time series as a non-closed recursive formula. Autoregressive modelling has been used before for time series analysis in asteroseismology (Kovacs 1983) although it has not become popular due to the difficulties for the physical interpretation of the fits, in contrast to the well-known Fourier frequencies.

We are interested here in the presence of any non-analytic content of the underlying function describing the observed signals. In contrast to the cubic splines, ARMA models are able to represent non-analytic functions, thereby allowing us to identify non-analytic signals from the residuals.

The algorithm used to calculate ARMA models is similar to gap-filling algorithm MIARMA (Pascual-Granado 2015). This is based on minimizing rms residuals and hence, is insensitive to the arrow of time (Scargle 1990). However, here we extrapolate only one point each time and the segments of the modelled data are short enough to be able to estimate derivatives without bias. The invariance to the arrow of time allowed us to calculate backward extrapolations simply as an inverted forward extrapolation.

Like cubic splines modelling, the accuracy of the extrapolations with ARMA models depends on the number of data points of the segment to be modelled. In addition, ARMA extrapolations depend on two parameters: the orders p and q , which are the number of autoregressive coefficients and the number of moving average coefficients, respectively. Hereafter, ARMA descriptions of the time series will be denoted by its coefficients as $\text{ARMA}(p, q)$. In general, an ARMA model is able to reproduce a sinus wave with only two terms. Accordingly, the number of terms increases with the frequency content of the signal. The ARMA modelling requires of an iterative process involving the following steps:

1. Identification of the order of the starting model using Akaike criterion (Akaike 1974) and described in detail in Pascual-Granado (2015). This affects principally the velocity of convergence but does not modify the final result.
2. Calculation of the parameters using a Steiglitz-McBride algorithm (Steiglitz 1965) given the orders for the AR and MA contributions.
3. Evaluation of the validity of the model based on the goodness of the fit in backward and forward extrapolations. The algorithm starts again with increased orders until a minimum in the residuals is reached.

4. Connectivity analysis applied to stellar light curves

The test was applied to the light curves of two delta Scuti stars observed by CoRoT and *Kepler* satellites: HD174936 and KIC006187665. The parameters required for this modelling are specific for each of the studied time series. However, in order to avoid misleading interpretation of spurious numerical effects we maintain the same parameters when characterising the connectivities for both time series. The δ Scuti star HD 174936 was taken as a reference.

For the first step of the ARMA approach, the initial tentative orders were obtained using autoregressive methods (AR).

Then we iterated until minimal randomly distributed residuals were obtained, fixing as optimal orders $p = 20$ and $q = 1$. Cubic splines and $\text{ARMA}(20,1)$ connectivities were then calculated for every point in a segment of 1000 samples of the time series using subsets of 80 data points (40 forward and 40 backward extrapolations). We adopted this number of samples per segment as a balance between computer time consuming and reliability of the ARMA modelling. In any case, we checked that the statistical properties of both approximations do not change significantly for a higher number of datapoints modelled.

In order to avoid artifacts we have selected data segments with no gaps and homogeneous sampling. Also, studied segments were statistically normalized allowing comparison of the connectivity results obtained using different time series.

4.1. HD 174936

We first applied the method to the light curve of the multi periodic δ Scuti star HD 174936 using as analytic solution the 422 independent frequencies found in its periodogram (García Hernández 2009).

These are distributed in a range of frequencies below $900 \mu\text{Hz}$. This very large range of excited frequencies is not predicted by any non-adiabatic model describing opacity-driven pulsators.

The numerical methods described above were applied to an analytic model of this star. The time series of this model was calculated using the amplitudes and phases of the frequencies found in García Hernández (2009) for the signal component, and a noise component based on additive gaussian white noise with a level obtained from the same reference.

For both the observed data (Fig. 1) and the time series of the analytic model (Fig. 2), the connectivities calculated using cubic splines have higher dispersion than the ARMA ones. Nevertheless, a striking effect is seen in the observed data that can not be noticed in the analytic model: spline connectivities are not randomly distributed but strongly correlated with the signal, suggesting that these data are not sufficiently well described by the analytic model. The same effect is present, although five times smaller in amplitude, for the ARMA results.

The difference between the analytic model and observed data has been usually interpreted as correlated noise caused by the effects of the turbulence present in the envelopes of stars (Chaplin 1997). In order to check if the convection is at the origin of this phenomenon, some additional tests were performed by adding non-white noise to the analytic model following the procedure described in Kallinger (2010). A power law with two components (36 ppm and 19 ppm of amplitudes) was added in the frequency domain to the analytic model (see Appendix B). The results of this test do not significantly change the previous ones, i.e. correlated connectivities were found. We conclude that the fully developed turbulence as represented by Harvey models (Harvey 1985) is not at the origin of this phenomenon.

4.2. KIC 006187665

In order to discard possible instrumental effects specifically linked to a given instrument, a similar test was applied to the time series from another star KIC 006187665 observed by the *Kepler* satellite, and classified as a hybrid Gdor/Dscuti pulsator by Uytterhoeven (2011). The standard Fourier analysis of the light curve supplied by the satellite at a sampling rate of 60 seconds (short cadence regime), yielded 659 significant peaks.

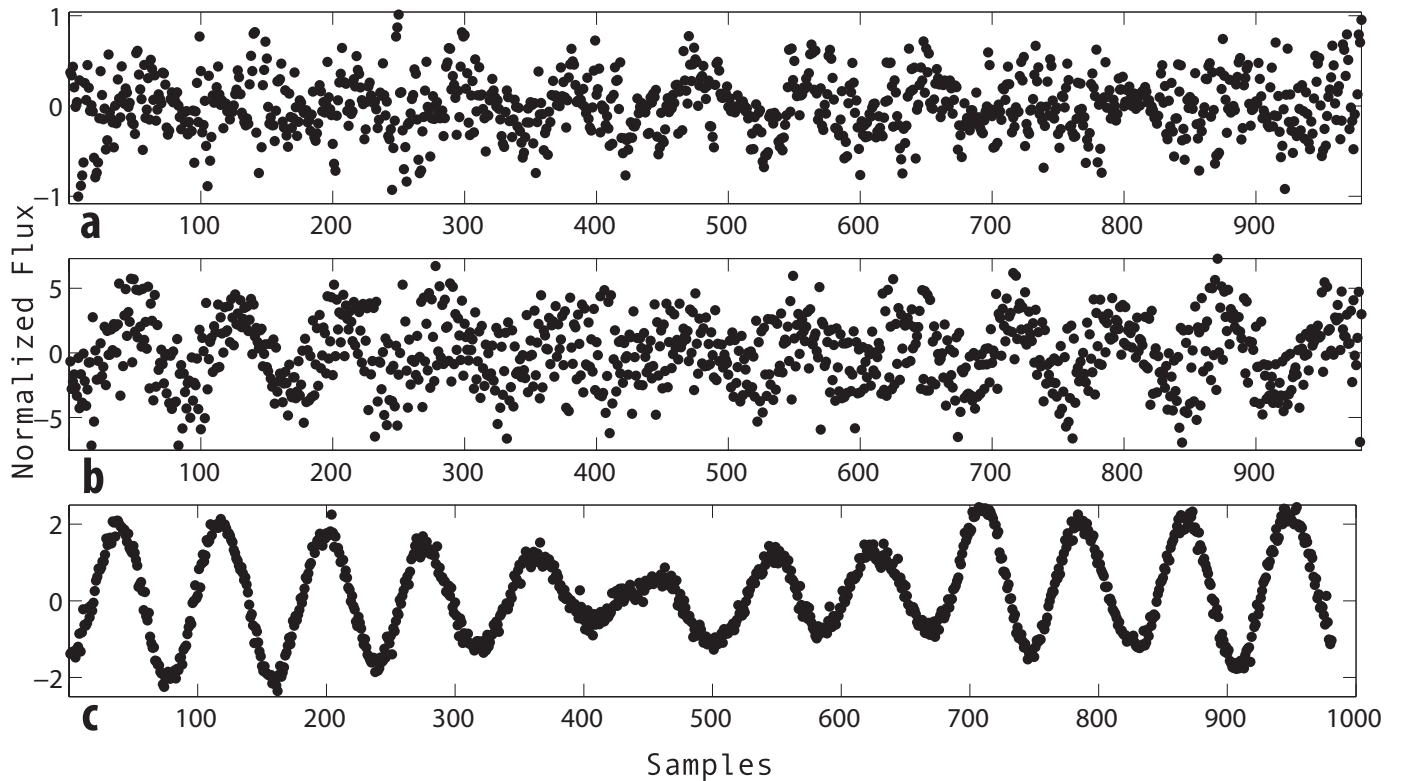


Fig. 1. Connectivities of the CoRoT data for the δ Scuti star HD 174936. **a**, ARMA connectivities. **b**, cubic splines connectivities. **c**, the original light curve. (Note the different scaling factors in top and middle panels.)

The connectivities corresponding to the *Kepler* target show a very similar behaviour to the CoRoT ones (Fig. 3a,b), that is, cubic splines connectivities show a more dispersed distribution than the ARMA connectivities, and strongly correlated with the original *Kepler* time series. We have calculated the Pearson correlation coefficients for both splines and ARMA connectivities of the two time series studied here confirming these results (see table C1).

The non-random structures of the connectivities reflect the non-analytic component of the studied function. In order to characterize them, we calculated their amplitude spectrums (Fig. C1) after verifying its applicability by using the test described in this paper (see Appendix D). The low frequency range in the case of splines connectivities is very similar to that found in the amplitude spectrum of the original time series. However a regular pattern can be seen along the frequency axis increasing in amplitude with frequencies. The ARMA connectivities show some pattern at low frequencies whereas an almost flat spectrum, appearing as a white noise spectrum, is obtained at higher frequencies. The amplitudes of the ARMA connectivities are 1-2 orders of magnitudes lower than those calculated with cubic splines.

5. Discussion and conclusions

The test introduced here through the connectivities allows us to fully characterise the differentiability of the underlying function of a time series. We remark that although the underlying function can be non-analytic due to the noise itself being non-analytic, our test assumes the non-analyticity of this component and evaluates the deterministic component through the correlation of the connectivities.

We have shown that at least for the stars studied here the functions underlying the observed time series are non-analytic. This comes from the finding that connectivities theoretically expected to be zero or a white noise distribution are strongly correlated with the signal. We have shown this to be an intrinsic property of those functions, which implies that their Fourier expansion convergence is not guaranteed. In addition, the fact that periodograms actually converge to a real value does not guarantee that this value is an asymptotically ($N \rightarrow \infty$) unbiased estimator of the spectral density of the underlying function. Therefore, the value estimated for the spectral density at a given frequency bin is not necessarily the true value.

In order to illustrate the possible relationship between the non-analyticity and the unexpected huge number of frequencies found in multi periodic stars, we performed a connectivity and frequency analysis of simulated data from a known function, the Weierstrass function (Bailey 2007). This function is originally defined by an infinite sum of harmonic components (see Appendix E) but can be represented using eq. E.2 with just a few components (Weisstein 2015). However, it is considered a pathological function since it is continuous everywhere but differentiable nowhere. We show that when performing frequency analysis of the simulated data the residuals of the prewhitening process does not converge to white noise. Notice the large number of frequencies obtained as compared with the few components used to generate the function (see Fig. E.3). This is an example of an inconsistency in the application of harmonic analysis that can be related to the non-analyticity of the underlying function.

We are not claiming that Weierstrass function is at the basis of the inconsistencies found in the time series of the two δ Scuti stars observed from space with different instrumentation, we are only using its pathological properties to try to "simulate"

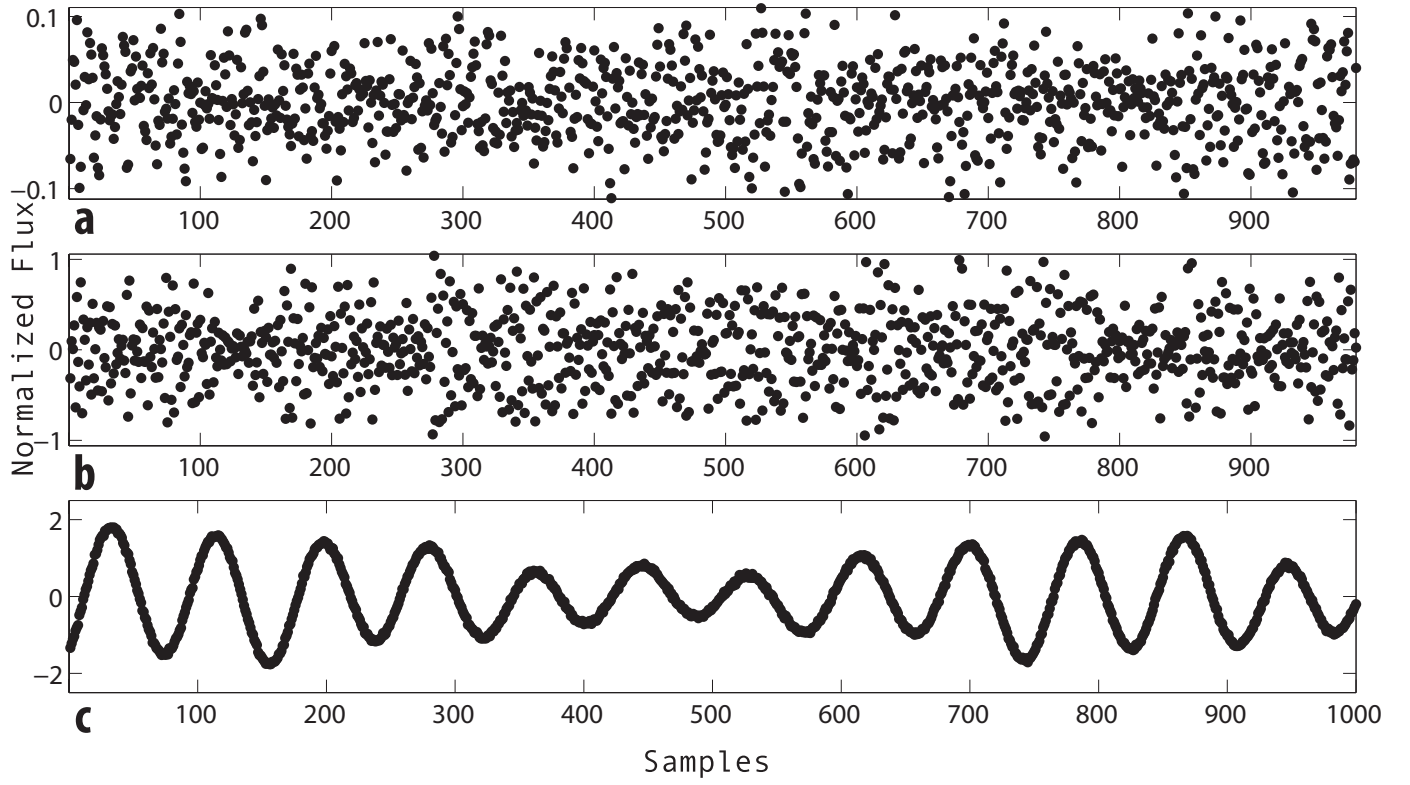


Fig. 2. Connectivities of an analytic model of the star HD 174936 based on the first 422 frequencies detected using standard Fourier techniques. **a**, ARMA connectivities. **b**, cubic splines connectivities. **c**, the original light curve. (Note the different scaling factors in top and middle panels.)

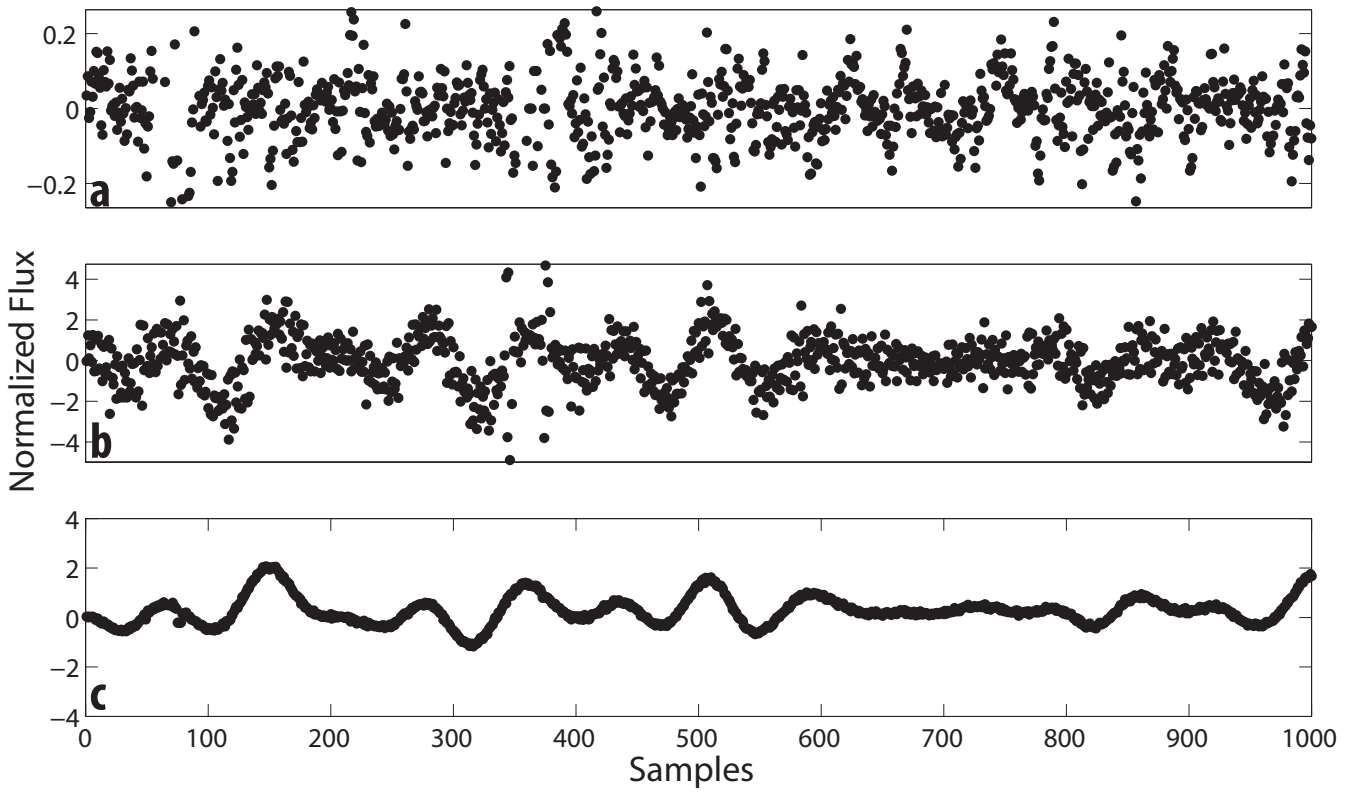


Fig. 3. Connectivities of the *Kepler* data for the hybrid star KIC 006187665 and the original light curve for comparison. **a**, ARMA connectivities. **b**, cubic splines connectivities. **c**, *Kepler* light curve for comparison. (Note the different scaling factors in top and middle panels.)

the real cases here studied. Although this peculiar function suggests a Fourier expansion in its standard form (eq. E.1) Fourier frequencies, given by an integer number of the argument, cannot be defined. This could point towards a reinterpretation of the term frequency in the context of the harmonic analysis of pulsating stars. The consequences of this new description of frequency are obviously out of the scope of this paper. Although the origin of these inconsistencies is unclear, our next step is to try to correct numerically the time series in order to overcome the analyticity test and therefore obtain a cleaner power spectrum.

We conclude that periodograms do not provide a mathematically consistent estimator of the frequency content for stellar variability of the objects here studied. Thereby, the concept of unambiguous detection of a physical frequency should be revised. This constitutes the first counterexample against the current paradigm which considers that any physical process is described by a continuous (band-limited) function that is smooth and infinitely differentiable.

This could be on the basis of the unknown nature of many phenomena related to the power spectrum of AF pulsating stars: the unexpected huge number of frequencies found in multi periodic stars and the range of these frequencies as mentioned before, and also the ubiquitous presence of correlated noise in the residuals of the fitting of the light curves.

This paper shows an inconsistency in the application of harmonic analysis to some pulsating stars observed by space missions like CoRoT or *Kepler*. This inconsistency is related with the non-analyticity of the underlying function, the origin of which is still to be unveiled.

Acknowledgements. The CoRoT space mission, launched on December 27th 2006, has been developed and is operated by CNES, with the contribution of Austria, Belgium, Brazil, ESA (RSSD and Science Programme), Germany and Spain. Funding for the *Kepler* Discovery mission is provided by NASA's Science Mission Directorate. The authors acknowledge support from MINECO and FEDER funds through the Astronomy and Astrophysics National Plan under number AYA2012-39346-C02-01. J.P.-G. acknowledges support from MINECO through the FPI grant number BES-2008-008252. JCS acknowledges support by the European project SpaceInn, with reference 312844 within the European SPACE program FP7-SPACE-2011-1, and also acknowledges funding support from the Spanish "Ministerio de Economía y Competitividad" under "Ramón y Cajal" subprogram.

Appendix A: Kolmogorov continuity theorem

As mentioned in Sect.3 this theorem (Revuz 1999) allow us to fully characterize the properties of a stochastic function through its continuous representation. The theorem comes after the following definition:

Def.: An extension (\tilde{Y}_t) of a stochastic process (Y_t) is a process such that for every $t \geq 0$, $P(Y_t = \tilde{Y}_t) = 1$. As a corollary, if the process (Y_t) satisfies for all times $T > 0$ and $0 \leq s, t \leq T$,

$$E(\|Y_t - Y_s\|^\alpha) \leq c |t - s|^{1+\varepsilon}, \quad (\text{A.1})$$

with α, ε, c positive constants, then there exists an extension of the process (Y_t) that is a continuous process whose paths are almost surely continuous. This is known as topological separability.

The conditions above mentioned are quite mild and they allow us to consider stochastic processes as continuous paths instead of discrete values as it is the case for deterministic functions.

It can be demonstrated that these conditions are fulfilled by any white noise process and even for a non-gaussian noise hav-

ing a spectral density proportional to f^α (with $\alpha < 0$), i.e. the so-called coloured noise.

As an example, a real-valued Wiener process (so-called coloured noise proportional to ν^{-2}) has continuous trajectories a.s.⁵ by this theorem. This means that it admits a separable and progressively measurable extension (Capasso 2005).

In any case, in this work we refer to the stochastic component of a time series with the properties of gaussian white noise.

Appendix B: The analytic model construction

The numerical methods described in the main text were applied to an analytic model of the δ Scuti HD 174936. This model was built using the amplitudes and phases of the first 422 frequencies found using a standard Fourier analysis in García Hernández (2009), fitted to an harmonic Fourier-like expansion of the form

$$x^a(t_i) = \sum_{i=1}^n A_i \cos(2\pi \nu_j t_i + \phi_j) + N^w + N^{nw} \quad (\text{B.1})$$

where A_i , ν_j , and ϕ_j represent respectively the observed amplitudes, frequencies, and phases. The $N^w + N^c$ terms represent the white noise, defined as normalised distribution scaled to the data,

$$N^w = \langle \sigma^N \rangle \quad (\text{B.2})$$

and the non-white noise, calculated as

$$N^{nw} = \mathcal{F}^{-1}[\tilde{N}^{nw}] \quad (\text{B.3})$$

where \tilde{N}^{nw} is defined in the transformed domain as

$$\tilde{N}^{nw} = \left[\sum_{k=1}^m \frac{a_k}{1 + \left(\frac{\nu}{b_k}\right)^{c_k}} \right]^{1/2} e^{i\theta} \quad (\text{B.4})$$

where a_k and b_k are free parameters. The values of $c_k = 4$ and $m = 2$ were adopted following Kallinger & Matthews (2010). This noise was calculated setting the θ phases equal to a random uniform distribution in the $[-\pi, \pi]$ interval. The light curve so constructed is supposed to be originated by the oscillation modes of the pulsating star plus granulation noise modelled as a fully developed turbulence regime (Harvey 1985) and given by eq. B4.

Appendix C: Pearson correlation coefficients

The correlations between the connectivities and their corresponding original time series are evident in Fig.C.1, even so, we calculated their corresponding standard Pearson coefficient in order to quantify the correlation between these two series (see Table C.1). We also calculated P-values for testing the null hypothesis (i.e. no correlation). P-values were calculated using a t-Student distribution. Considering the P_s negligible values obtained for the time series of the selected stars, the null hypothesis must be rejected under any computable limit of validity, i.e. the two series are strongly correlated. On the other hand, for the case of ARMA (P_a), this limit can be put as an standard choice for this parameter, e.g. $P < 0.05$, i.e. only in 5% of the cases the ARMA connectivities would be correlated with the original signal.

Since ARMA models can capture the non-analytic component of the signal it is expected such non-correlation of the connectivities.

⁵ almost surely in mathematical jargon

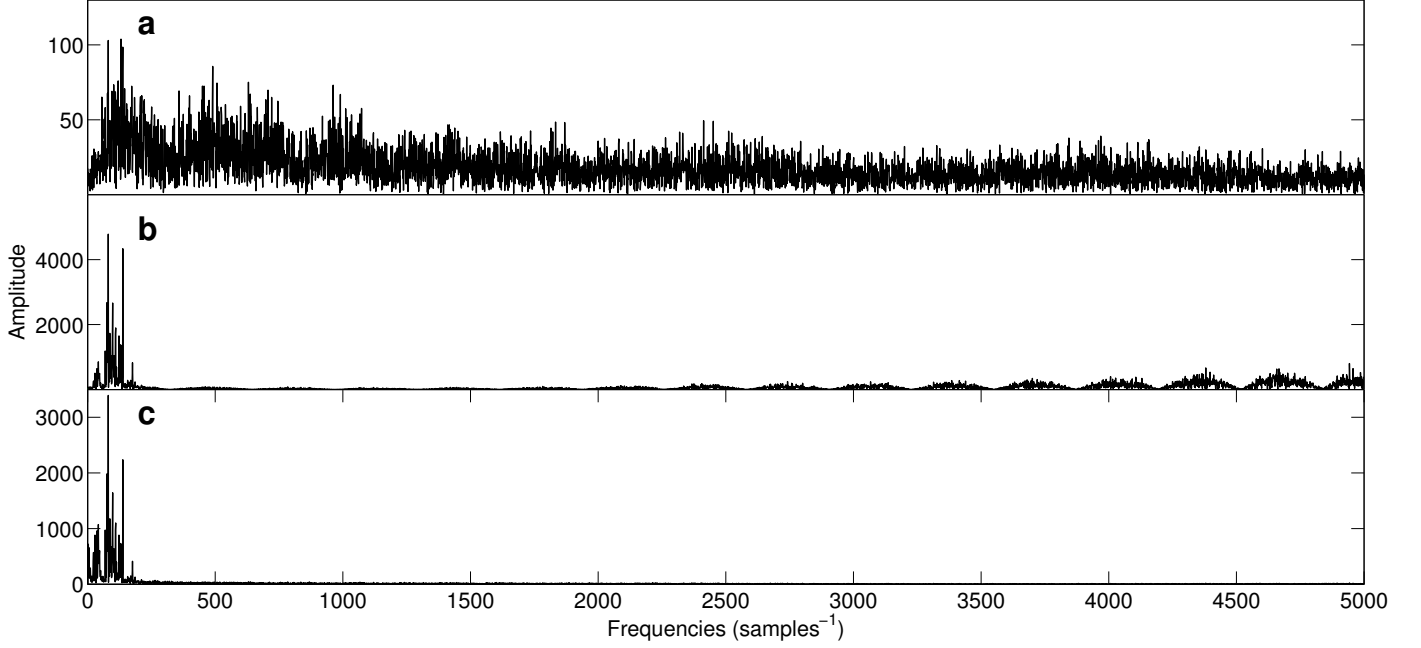


Fig. C.1. Amplitude spectrum of the connectivities of the *Kepler* data for the hybrid star KIC 006187665. From top to bottom: panel (a) shows the amplitude spectrum of the ARMA connectivities, panel (b) amplitude spectrum of the cubic splines connectivities, panel (c) amplitude spectrum of the original light curve for comparison.

Table C.1. Pearson correlation coefficients (ρ) and the associated probabilities of no-correlation (P) for ARMA and splines connectivities for the time series considered in this work. Subscripts a and s refer to the ARMA and splines approaches, respectively.

Time series	ρ_a	P_a	ρ_s	P_s
CoRoT 7613	-0.0075	0.8139	0.4243	5.8247e-45
Analytic model	0.0072	0.8199	8.371e-4	0.9789
KIC 006187655	-0.0257	0.4176	0.5553	5.8039e-82

Appendix D: Second-order connectivities

Analyticity, understood as infinite differentiability, guarantees the convergence of the Fourier expansion of a function, so allowing to perform Fourier analysis. The test described in this work indicates whether the function underlying a time series is differentiable or not. Considering the connectivities as a new time series, the calculation of the connectivities of the connectivities (from now on second-order connectivities) informs us about their differentiability. If, using Pearson correlation coefficients, second-order connectivities are independent, normal, and randomly-distributed, then connectivities are differentiable. It is thus legitimate to perform a Fourier analysis of them.

We applied this method presented to the splines connectivities of the stellar light curve of KIC 006187665 (Fig. D1, panel a). Contrary to the connectivities, second-order connectivities calculated using the splines approximation (Fig. D1, panel b) show a normal white noise distribution. These, when calculated with splines show that the connectivities are differentiable. The ARMA approximation does not provide better fits than the cubic splines in this case.

Appendix E: Non-analytic model

Frequency and connectivity analysis has been applied to a non-analytic model based on a numerical realization⁶ of the Weierstrass function. This is an extreme case of a function exhibiting non-analyticity because it is continuous everywhere but differentiable nowhere.

Appendix E.1: Definition

The original definition of the Weierstrass function is:

$$W(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x) \quad 0 < a < 1, \quad a \cdot b > 1 + \frac{3}{2}\pi \quad (\text{E.1})$$

and b , a positive odd integer number. For rational numbers $x = \frac{p}{q}$, which is our case, $W(x)$ can be calculated with a finite sum such as:

$$W\left(\frac{p}{q}\right) = \frac{\pi}{4q^2} \sum_{n=1}^{q-1} \frac{\sin\left(\frac{n^2 p \pi}{q}\right)}{\sin^2\left(\frac{n \pi}{2q}\right)} \quad (\text{E.2})$$

This formula allows to generate an exact numerical realization of the Weierstrass function without calculating an infinite sum.

For the non-analytic model that we used we generated three realizations of $W(x)$ with different q values and obtained a 30000 datapoints time series by iterating on p and adding each q component.

Appendix E.2: Connectivity analysis

We apply the numerical techniques introduced before to the time series generated as described in previous section. Connectivities

⁶ obtained by sampling the continuous function

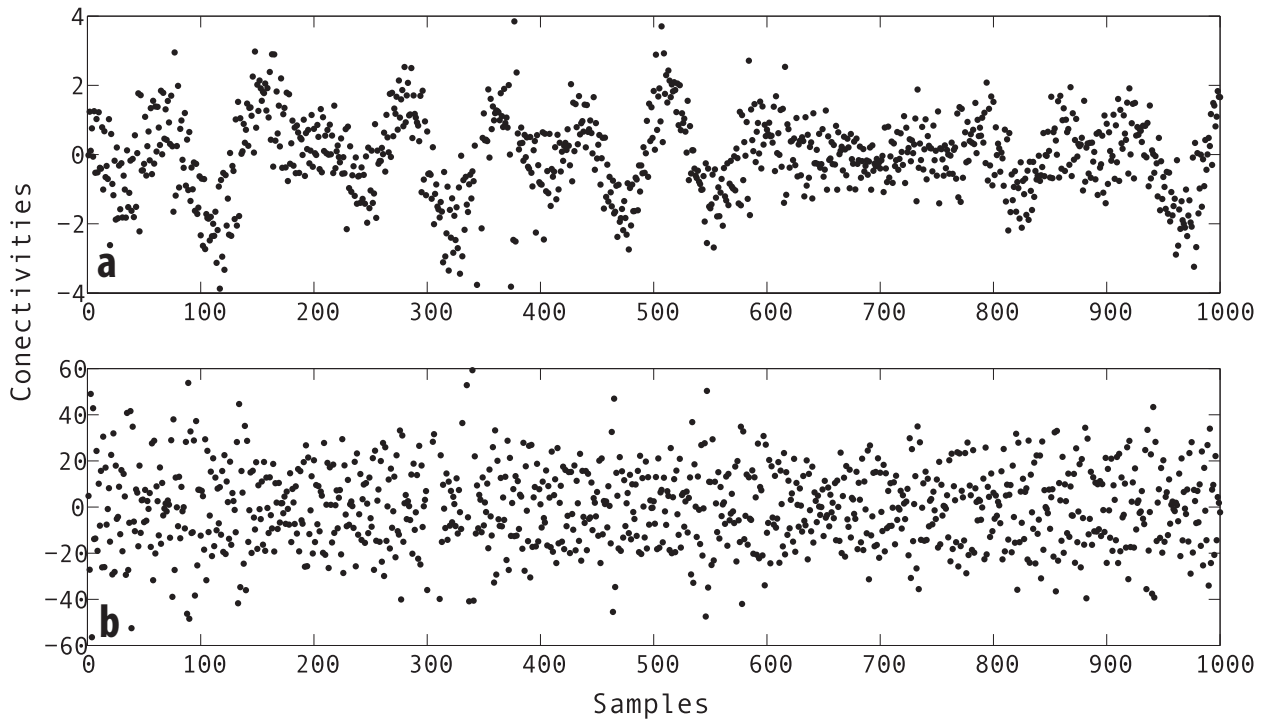


Fig. D.1. Second order connectivities calculated using the cubic splines approximation for *Kepler* data of the the hybrid star KIC 006187665, and the first order connectivities for comparison. From top to bottom: Panel (a) first order spline connectivities, panel (b) second order splines connectivities. Note that the correlation present in the first is lost in the last.

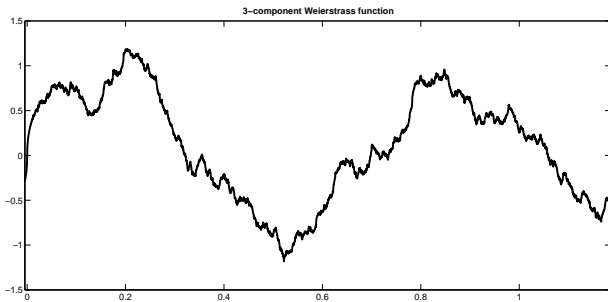


Fig. E.1. A realization of the Weierstrass function with three components corresponding to three different q values as indicated in eq.E.2.

calculated using the splines approximation are clearly correlated (Fig.E.2) and ARMA connectivities are not. These results are consistent with the definition and properties of the Weierstrass function and demonstrate with the analytic model discussed in Appendix B the validity of the connectivity analysis as a test for the analyticity of the underlying function of a time series.

Appendix E.3: Frequency analysis

A power spectrum of the time series of the non-analytic model and frequency detection was performed using *SIGSPEC* (Reegen 2007) in order to check the effect of the non-analyticity in a standard Fourier analysis. A sequence of power spectra was obtained during the prewhitening process of different sets of frequencies (Fig.E.3). The power spectrum at each step was significantly different to a white noise spectrum and the prewhitening process continued until the program stopped at 738 frequencies when the limit in significance was reached.

In conclusion, a frequency detection procedure using standard Fourier analysis is not convergent. This is due to the fact that, in this case, the power spectrum calculated with a DFT is not a consistent estimator for the original frequency content as a consequence of the non-analyticity of the underlying function.

References

- Akaike, H., 1974, IEEE Trans. Automatic Control, AC-19, 716-723
- Auvergne, M., et al. 2009, A&A, 506, 411
- Bailey, D. H., 2007, Experimental mathematics in action, Ak Peters Series
- Boor, Carl de, 1978, A Practical Guide to Splines, 1st ed., Series: Applied Mathematical Sciences, New York Berlin Heidelberg: Springer, 27
- Box, George E. P., Jenkins, & Gwilym 1976, M. Time Series Analysis: Forecasting and Control. San Francisco, CA. Holden-Day
- Burden, R. L. & Faires, J. D. 2011, Numerical Analysis, 9th ed. Boston, MA: Brooks/Cole
- Capasso, V., Bakstein, D., 2005, An Introduction to Continuous-Time Stochastic Processes, Birkhäuser
- Carleson, L. 1966, On convergence and growth of partial sums of Fourier series, Acta Math., 116(1), 135
- Chaplin, W. J. et al. 1997, MNRAS, 287, 51-56
- Diestel, R., Graph Theory, Electronic ed. 2005, p.12
- Dijk, G. van 2009, Introduction to harmonic analysis and generalized Gelfand pairs, Berlin, New York. Walter De Gruyter
- García Hernández, A. et al. 2009, A&A, 506, 79
- Gilliland, R. L. et al., 2010, Publ. Astron. Soc. Pac., 122, 131
- Harvey, J. 1985, Proc. ESA Workshop on Future Missions in Solar, Heliospheric & Space Plasma Physics, Garmisch-Partenkirchen, 199
- Jech, T. J. 1997, Set Theory, 2nd ed. Berlin: Springer-Verlag
- Kallinger, T. & Matthews, J. M. 2010, ApJ, 711, L35-L39
- Kovacs, G. 1983, Solar Physics, 83, 123-128
- Kaplan, W. 1992, Advanced Calculus, 4th ed. Reading, MA: Addison-Wesley
- Pascual-Granado, J., Garrido, R., Suárez, J. C., 2015, A&A, 575, A78
- Poretti, E., Michel, E., Garrido, R., et al. 2009, A&A, 506, 85
- Revuz, D. & Yor, M. 1999, Continuous Martingales and Brownian Motion, A Series of Comprehensive Studies in Mathematics (Springer), 293
- Reegen, P., 2007, A&A, 467, 1353
- Royden, H. L. 1988, Real Analysis, Prentice-Hall
- Scargle, J. D. 1990, AJ, 359, 469 (1990).

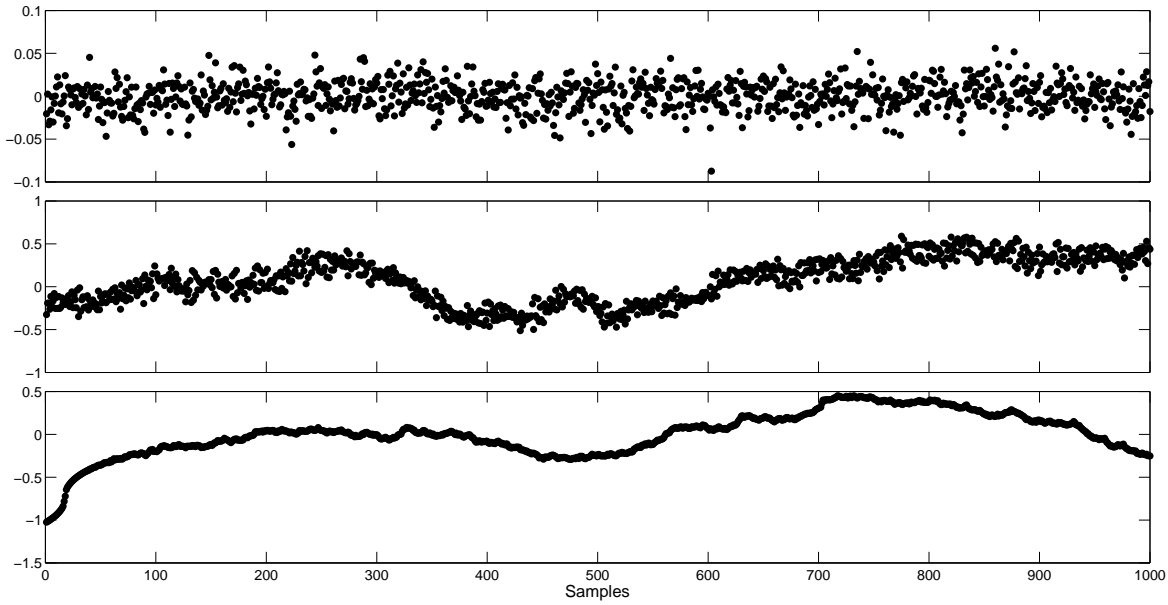


Fig. E.2. Connectivities of a realization of a 3-component Weierstrass function. **Upper panel:** ARMA connectivities. **Middle panel:** cubic splines connectivities. **Lower panel:** the original light curve. (Note the different scaling factors.)

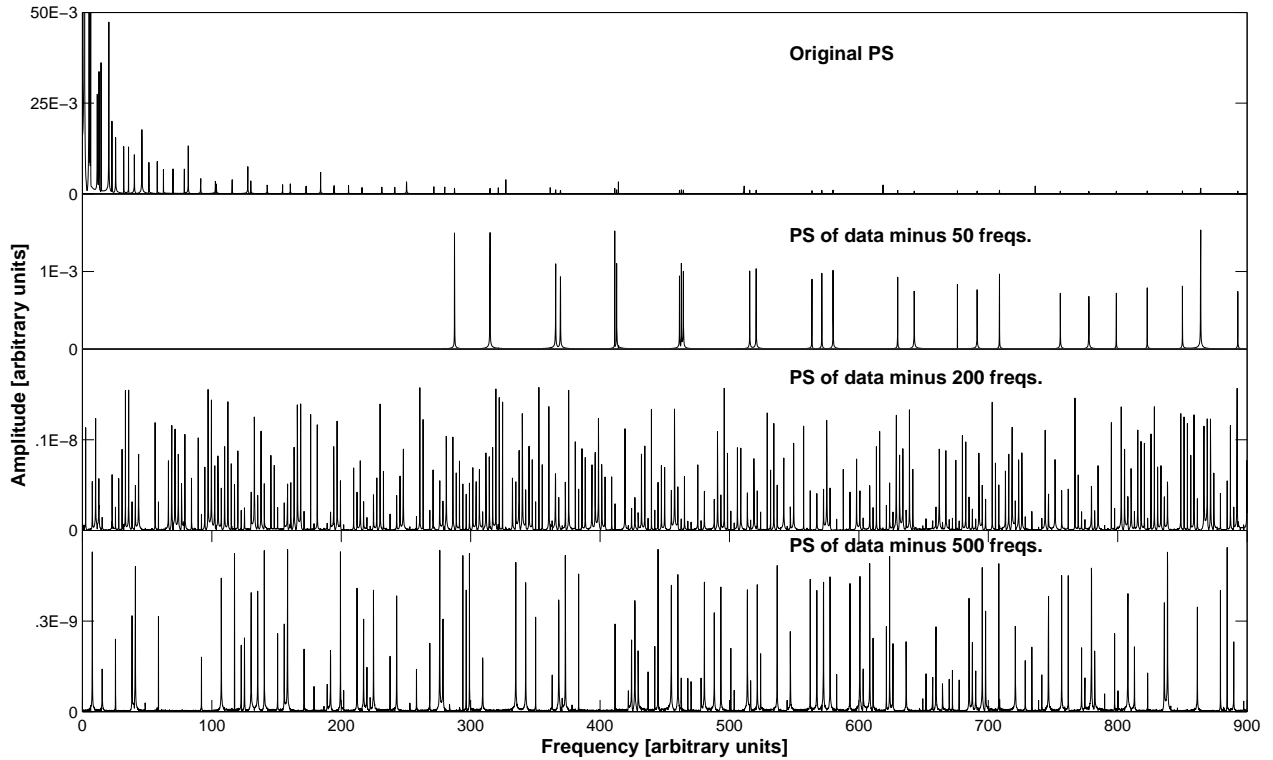


Fig. E.3. Power spectrum of the time series generated with a realization of a 3-component Weierstrass function (top panel) and the power spectra obtained during different steps in the prewhitening process (successive panels).

- C. E. Shannon 1949, Communication in the presence of noise, Proc. IEEE, 37, 10
- Steiglitz, K., & L.E. McBride. 1965, IEEE Trans. Automatic Control AC-10, 461
- Weisstein, Eric W. "Weierstrass Function." From MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/WeierstrassFunction.html>
- Wiener, N. 1923, Differential Space. J. Math. and Phys., 2, 131.
- Wold, H. 1938, A Study in the Analysis of Stationary Time Series 2nd ed., Uppsala. Almqvist and Wiksell
- Uytterhoeven, K. et al. 2011, A&A, 534, A125